

## University of Groningen

### Ising models and neural networks

Schaap, Hendrikjan Gerrit

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

2005

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Schaap, H. G. (2005). *Ising models and neural networks*. s.n.

#### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

#### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

# Samenvatting

Denk eens aan een rivier, voortdurend stromende volgt hij zijn weg. De aard van de stroom verandert niet: het water van de rivier volgt altijd de weg van de minste weerstand. Stel je<sup>3</sup> eens voor dat het hard regent, het waterpeil van de rivier wordt hoger en hoger. Dan, opeens, is het water zo hoog gestegen dat de oevers overstromen. Daardoor kan het water in de nabijgelegen weilanden vloeien. Het zal nu wel even duren voordat de rivier zich heeft aangepast aan deze nieuwe situatie.

In de statistische fysica proberen we natuurlijke processen te modelleren. In het bijzonder proberen we de verbinding te leggen tussen de micro- en macroscopische eigenschappen. Voor de rivier koppelen we de globale stroming aan de beweging van de watermoleculen. Sommige universele wetten, betreffende enige globale (= macroscopische) kenmerken, beschrijven de algemene eigenschappen van de rivier, dit ondanks dat er een enorme hoeveelheid aan watermoleculen bij betrokken is. Daartoe gebruiken we kanstheoretische methoden (stochastiek) nodig om toe te passen op de onderliggende microscopische differentiaalvergelijkingen waaruit de beweging van de watermoleculen volgt. Als we dan kijken naar de macroscopische eigenschappen, weten we het globale gedrag van de rivier. Dit globale gedrag kan worden beschreven door vergelijkingen die alleen afhankelijk zijn van macroscopische eigenschappen. Het onderliggende microscopische deel is verdwenen door onze toegepaste stochastiek.

We pakken eens een dobbelsteen waarmee we een paar van de betrokken stochastische principes gaan demonstreren. Zoals we weten, hebben we dezelfde kans om een 1 of een 6 te gooien. Maar door ervaring weten we dat na een klein aantal worpen, de ontstane *relatieve frequenties* van de gegooide cijfers, significant van elkaar kunnen verschillen. Alleen als we een dobbelsteen heel vaak gooien, zullen de ontstane relatieve frequenties meer en meer gelijk worden. Hetzelfde resultaat verkrijgen we wanneer we in plaats van één dobbelsteen heel vaak, een heleboel dobbelstenen tegelijk gooien en dan pas te kijken naar de relatieve cijferfrequentie. Natuurlijk, het is evident dat met een dobbelsteen 1000 keer gooien, hetzelfde is als met 1000 dobbelstenen één keer gooien. Uiteindelijk zal de relatieve frequentie van elk van de zes cijfers  $\frac{1}{6}$  naderen: gemiddeld komt elk cijfer een keer voorbij als we een dobbelsteen zes keer gooien. Deze relatieve frequentie wordt soms ook wel de *kans* op het cijfer genoemd. Met een kortere notatie noteren we dit als  $P(i)$ . Voor elk cijfer  $i$  op onze dobbelsteen geldt dat  $P(i) = \frac{1}{6}$ . De functie die aan elk cijfer  $i$  de bijbehorende kans toekent, noemen we de *waarschijnlijkheidsverdeling* van de eigenschap.

Al de materie om ons heen bestaat uit atomen. Elke gram van materie bevat zo

---

<sup>3</sup>Ik heb bewust gekozen voor een informeel getinte samenvatting. Wanneer U niet wenst getutoyeerd te worden dan mag U in plaats van je U lezen.

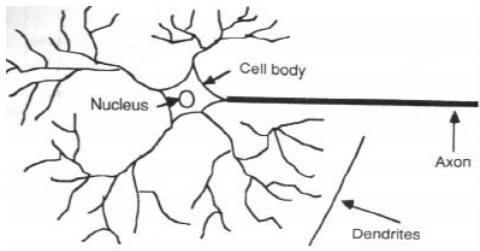
rond de  $10^{23}$  atomen. Vaak kijkt men naar een collectie van een aantal globale (bulk) eigenschappen als toevoeging bij de atomaire eigenschappen. In de beschrijving van de materie beschrijven alle atomen bij elkaar met de bovengenoemde globale eigenschappen het *systeem*. Elke bepaalde realisatie van de bijbehorende atomaire waarden noemen we de *configuratie* van het systeem. Het bepalen van de configuratie heeft wel wat weg van het in één keer gooien van  $10^{23}$  dobbelstenen.

Stel je nu voor dat je een macroscopische eigenschap wil meten, bijvoorbeeld de gemiddelde dichtheid. Omdat het aantal atomen heel groot is, is er geen noodzaak om in de waarschijnlijkheidsverdeling van de atomaire waarden de posities van alle afzonderlijke atomen bij te houden. Bij de rivier is de tijd die de globale aanpassingen in beslag nemen extreem lang, als je het vergelijkt met de tijdschaal van de lokale beweging van de watermoleculen. De manier waarop de eigenschappen van het systeem evolueren, noemen we de *dynamica* van het systeem. In het globale gedrag van de stroming van de rivier zijn de microscopische bewegingen van de watermoleculen uitgemiddeld.

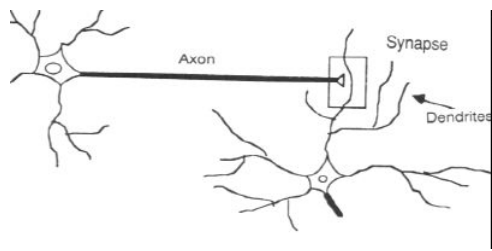
## Neurale netwerken

Het eerste onderwerp van het proefschrift is over een model dat zijn oorsprong heeft in de theorie van neurale netwerken. In het bijzonder willen we graag het begrip geheugen beter begrijpen. Onze hersenen zijn opgebouwd uit miljoenen neuronen die op een zeer complexe manier met elkaar verbonden zijn. Deze structuur noemen we een *neuraal netwerk*. Het is moeilijk om rechtstreeks dit netwerk te bestuderen, omdat er in een relatief klein gebied erg veel neuronen betrokken zijn. Om te begrijpen hoe het geheugen werkt, is een gebruikelijke methode om een simpeler model te construeren dat nog steeds de hoofdkenmerken van het geheugen bevat. Net zoals het neurale netwerk van de hersenen moet het model voldoende robuust zijn: als er tussen neuronen signalen worden doorgestuurd dan kunnen er altijd kleine afwijkingen ontstaan. De hersenen zijn in staat om van het licht vervormde signaal de ruis te verwijderen en zo het pure signaal te reconstrueren. Voor een goed globaal overzicht van neurale netwerken zie [19]. Een neuron is opgebouwd uit drie delen: het cellichaam, de dendrieten en het axon, zie Figuur A.1. De dendrieten hebben een boomachtige vertakte structuur en zijn verbonden aan het cellichaam. Het axon is de enige uitgaande verbinding van het neuron. Aan het einde van het axon vertakt het zich en is het verbonden aan de dendrieten van de andere neuronen via synapsen. Tussen het uiteinde van elke tak en dendriet zit een smalle ruimte: de synapsspleet.

Neuronen communiceren met andere neuronen via elektrische signalen. De weg die het elektrische signaal aflegt van een neuron  $i$  naar een neuron  $j$  is als volgt, zie Figuur A.2. Eerst stroomt het signaal vanuit het cellichaam van neuron  $i$  in het axon dat verbonden is met neuron  $j$ . Dit is het *output*-signaal van neuron  $i$ . Als het signaal



**Figuur A.1:** Componenten van een neuron (overgenomen uit [19])



**Figuur A.2:** De synaps (overgenomen uit [19])

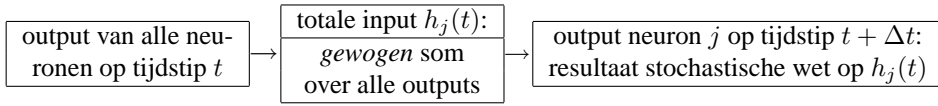
bij het einde van het axon komt, dan geeft het axon neurotransmitters af aan de synapspleet. Dan worden de neurotransmitters door de receptoren aan de dendriet van neuron  $j$  terugvertaald naar een elektrisch signaal. Er zijn verscheidene soorten neurotransmitters. Sommige versterken het binnenkomende signaal voordat het naar de dendriten gaat van andere neuronen, en andere verzwakken het.

Dit uiteindelijke signaal dat door de receptoren van neuron  $j$  geproduceerd wordt, noemen we het *input*-signaal van neuron  $i$  naar neuron  $j$ . Uiteindelijk komt het signaal aan in het cellichaam van neuron  $j$ . In het cellichaam van neuron  $j$  komen alle inputs bij elkaar. De cel verwerkt de inputs (die we wiskundig modelleren door het uitvoeren van een *gewogen som*). Dat noemen we de *totale input*  $h_j$  van neuron  $j$ . Afhankelijk van de uitkomst, zendt het cellichaam een nieuw signaal uit dat gaat naar het axon van neuron  $j$ , zodat het doorgestuurd kan worden naar andere neuronen. Dit noemen we de output of de *toestand* van neuron  $j$ .

Om een bruikbaar model te maken dat gebaseerd is op deze neurale processen is het nodig om enige flinke vereenvoudigingen door te voeren. Als eerste vereenvoudiging nemen we aan dat elk neuron interactie heeft met elk ander neuron: het neurale netwerk is *volledig verbonden*. Verder nemen we aan dat elk neuron slechts twee mogelijke outputs heeft. We schrijven  $\sigma_i$  als we de toestand van neuron  $i$  bedoelen:  $\sigma_i = +1$  als het *geëxciteerd* is en  $\sigma_i = -1$  als het zich in de *rusttoestand* bevindt. Ook nemen we aan dat het signaal niet verandert wanneer het door de synapspleet gaat. Hierdoor is de input naar neuron  $j$  die komt van neuron  $i$  gelijk aan de output  $\sigma_i$  van neuron  $i$  die gaat naar neuron  $j$ .

Om de dynamica van ons model te modelleren, introduceren we de tijd  $t$ . Na elke tijdperiode van  $\Delta t$  (met  $\Delta$  zeer klein) veranderen we tegelijkertijd de output van elk neuron. Het verwerkingsproces van het cellichaam modelleren we door twee stappen:

1. Op tijdstip  $t$  vermenigvuldigen we elke binnenkomende input van de andere neuronen met een gewicht. Hierover nemen we de som over alle neuronen (behalve neuron  $j$ ). Het resultaat is de totale input  $h_j(t)$  op tijdstip  $t$ .
2. De output  $\sigma_j(t + \Delta t)$  van neuron  $j$  op tijdstip  $t + \Delta t$  is de uitkomst van een waar-



**Figuur A.3:** *Dynamica van het neurale netwerk model*

schijnlijkheidsverdeling over de twee mogelijke neuron toestanden. Deze verdeling komt tot stand met behulp van een stochastische wet op  $h_j(t)$ .

We nemen aan dat de cellichamen van de neuronen de verbindingen op een symmetrische manier behandelen: het gewicht dat neuron  $j$  aan de input van neuron  $i$  toekent, is gelijk aan het gewicht dat neuron  $i$  aan de input van neuron  $j$  toekent. In realistische neurale netwerken geldt deze interactiesymmetrie in het algemeen niet. Figuur A.3 vat de dynamica van ons model nog eens samen.

De stabiele configuraties onder deze dynamica vormen het *geheugen* van het systeem. Stabiel betekent dat als je begint met een stabiele configuratie dat het systeem alleen in configuraties terecht komt die er heel erg veel op lijken (wat neuronconfiguraties betreft). Door het kiezen van geschikte gewichten kunnen we de dynamica zo afstellen dat het geheugen wordt gevormd door een eindig aantal  $m$  vooraf geselecteerde neuron configuraties  $\xi^{(m)}$ , die we ook wel *patronen* noemen.

De stochastische wet hangt af van een parameter  $\beta$ . De inverse  $T = 1/\beta$  hiervan heet *temperatuur*. Als parameter  $\beta$  groot is, dan heeft het neuron een sterke neiging (een grote kans) om gelijk te worden aan het teken van zijn totale input. Wanneer  $\beta$  oneindig nadert, dan verandert de stochastische wet in een deterministische. Als we dan een configuratie nemen die dicht genoeg bij bijvoorbeeld  $\xi^{(1)}$  ligt, dan evolueert het systeem naar configuraties die gelijk zijn aan het patroon  $\xi^{(1)}$ . Met andere woorden: het neurale netwerk *herinnert* zich de configuratie  $\xi^{(1)}$  vanuit zijn geheugen. Dit betekent dat de neuron configuratie gelijk wordt aan  $\xi^{(1)}$  en dat daarna het systeem in deze configuratie blijft, zie Sectie 2.3.2.

Om de capaciteit van het geheugen te vergroten, kun je natuurlijk de generalisatie maken naar een hoger aantal  $q$  van mogelijke neuron toestanden. Maken we in bovenstaand model deze generalisatie naar  $q > 2$  dan noemen we het ontstane model het Potts-Hopfield model.

In het model dat we in Hoofdstuk 3 hebben behandeld, kiezen we de gewichten in de totale inputs op een andere manier. Hiervoor moeten we eerst voor elk neuron een verzameling van  $p$  continue variabelen  $\xi_i^{(p)}$  definiëren: de patronen. We nemen een willekeurig gekozen realisatie van deze variabelen. Ze hebben een Gaussische verdeling. Met de waarden van de geïntroduceerde patronen  $\xi_i^{(j)}$  stellen we de gewichten voor de totale input vast. De gewichten van de totale input bepalen we door twee van deze Gaussische patronen. Het aantal mogelijke toestanden per neuron zetten we vast op drie.

Als we het aantal neuronen laten toenemen, dan gebeurt het volgende indien het aantal groot genoeg is: voor elk gegeven aantal van neuronen concentreert het geheugen zich rond zes neuron configuraties. Deze configuraties zijn aan elkaar gerelateerd door middel van een discrete symmetrie. Bij elke neuron configuratie hoort een punt in de macroscopische ruimte die wordt opgebouwd door enige macroscopische variabelen. De zes stabiele neuron configuraties zijn te verdelen in paren van lijnrecht tegenover elkaar staande punten. Als we het aantal neuronen laten toenemen, zie je altijd de discrete symmetrie terug. Als je echter kijkt naar de plaats van de zes neuron configuraties dan draaien ze rond op drie cirkels naarmate het aantal neuronen groter wordt. Dit is nader uitgewerkt in Hoofdstuk 3. In de rij van groeiende neuron aantallen vullen de optredende neuron configuraties met elkaar op een regelmatige uniforme wijze de drie cirkels op in de eerdergenoemde macroscopische ruimte.

## Ferromagneten

Het andere onderwerp van het proefschrift betreft een beroemd model voor magnetische materialen: het Ising model. In het algemeen zijn er vele soorten magnetisme. De zogenaamde *paramagneten* zijn metalen die alleen gemagnetiseerd zijn, wanneer we een extern veld aanbrengen. Een ander belangrijk type metaal zijn de *ferromagneten*. Deze metalen behouden hun magnetisatie, zodra ze blootgesteld zijn aan een magnetisch veld. Als we het metaal verwarmen, dan verdwijnt uiteindelijk dit effect: het metaal gedraagt zich als een paramagneet. Voor meer uitgebreide informatie over magnetisme zie [46].

## Ising modellen

Het Ising model is een model voor een ferromagneet. Om dit aannemelijk te maken, hebben we een aantal aannames nodig. We nemen aan dat de niet gepaarde elektronen in de buitenste schil *gelokaliseerd* zijn: d.w.z. sterk gebonden aan de bijbehorende atomen. Alleen deze ongepaarde elektronen zijn verantwoordelijk voor het magnetisme. In het Ising model nemen we aan dat voor elk atoom er maar een niet gepaarde elektron in de buitenste schil zich bevindt.

Elke elektron heeft een intrinsiek hoekmoment: *spin*. Deze spin is verantwoordelijk voor een magnetisch moment. Dankzij de quantummechanica kan de spin slechts twee oriëntaties t.o.v. dit magnetische moment hebben: *op* en *neer* [5]. Met enig misbruik van notatie noemen we deze oriëntaties de *waarden* van de spin. Vanwege de aanname over het aantal ongepaarde elektronen hebben we voor de totale spin per atoom ook maar twee oriëntaties. De meeste metalen hebben meer dan één ongepaard elektron in de buitenste schil. Deze metalen kunnen dan ook meer dan twee oriëntaties hebben voor de totale spin per atoom. De meeste vaste materialen hebben een kristalstructuur:

de atomen, of ionen of moleculen liggen in een regelmatig herhaald 3 dimensionaal patroon. Dit maakt een aantal spin oriëntaties qua energie voordelig.

In magnetiseerbare materialen is het metaal onderverdeeld in domeinen met een netto magnetische lading. De overgangsgebieden tussen deze domeinen noemen we *domain walls* [46]. Het Ising model heeft alleen configuraties waar twee naast elkaar gelegen elektronen gelijke of tegenovergestelde spins hebben. Als er een domain wall is, dan is de dikte ervan automatisch nul.

De interacties tussen de gelokaliseerde elektronen worden ook wel de *Weiss interacties* genoemd. In het algemeen komen twee types van interacties regelmatig voor: *naaste-buur* en *mean-field*. Voor mean-field interacties zijn de interacties tussen elk paar van atomen gelijk aan elkaar. In het Ising model beperken we ons tot naaste buur interacties. Dit houdt in dat we aannemen dat al de overige interacties tussen de elektronen nul zijn. Dit is een goede benadering voor de lanthaniden groep (een bepaalde groep van elementen). Ondanks dat het een simpel model is en voor andere magnetische metalen op zijn hoogst een ruwe benadering, is het een erg succesvol model. Het beschrijft het fenomeen van fase overgang (denk aan de overgang vloeibaar  $\rightarrow$  gas) en wordt gebruikt in allerlei praktische toepassingen. Bovendien is het exact oplosbaar in 1 en 2 dimensies.

Nu geven we een wiskundige beschrijving van het model. Pak een stuk roosterpapier. Elke punt waar een verticale lijn een horizontale lijn kruist noemen we een *site*. De horizontale en verticale lijnstukken die beginnen bij een site en stoppen bij de dichtstbijzijnde kruising noemen we *bonds*. Op elke site  $i$  bevindt zich een atoom dat een netto spin heeft, met twee mogelijke oriëntaties ten opzichte van de spins van de andere atomen die we aangeven met  $\sigma_i = +1$  (spin 'op') en  $\sigma_i = -1$  (spin 'neer'). Zowel het atoom als de bijbehorende spinwaarden noemen we *spin*. In ons geval is de spinconfiguratie  $\sigma$  een array, die de spinwaarden  $\sigma_i$  van elke site bevat.

Tussen elke naaste-buur spinpaar is er een *interactie*

$$E_{ij}^{\text{ui}} = -\beta\sigma_i\sigma_j \equiv J\sigma_i\sigma_j \quad (\text{A.65})$$

We noemen deze interactie ook wel de *interactie-energie* tussen de atomen op site  $i$  en  $j$ . De *energie* van een configuratie is het totaal van de interactie-energieën. De variabele  $\beta$  is afhankelijk van het type materiaal waar we naar aan het kijken zijn. Deze interacties bepalen de waarschijnlijkheden van de configuraties. In ferromagneten hebben de naaste-buur spins de neiging om gelijke waarden te hebben. Daarom kiezen we de interacties in het model zo, dat de kans groter is dat dit zo is: we nemen  $J < 0$ . Hoe hoger de energie des te lager de kans is op de configuratie. De kans op een configuratie is

$$P(\sigma) = \frac{\exp(\beta \sum_{i,j} \sigma_i \sigma_j)}{Z(\sigma)} = \frac{\exp(-\sum_{i,j} E_{ij}^{\text{ui}})}{Z(\sigma)} \quad (\text{A.66})$$

met  $Z(\sigma)$  de som van de noemer over alle configuraties. We zien dat naarmate de temperatuur daalt, de interactie (A.65) sterker wordt. Dan is het extra waarschijnlijk

dat naaste-buur spins dezelfde spinwaarde hebben. Met (A.66) zien we meteen dat voor  $T = 0$  slechts twee configuraties met minimale energie positieve kans hebben: elke spin heeft dezelfde waarde. Voor zeer hoge temperatuur heeft elke configuratie gelijke kans. Dan gedraagt het model zich als een paramagneet. De temperatuur is dus een maat van wanorde in het systeem. Voor lage temperaturen hebben de meeste spins gelijke waarden ten opzichte elkaar, voor hoge temperaturen zijn de waarden min of meer willekeurig verdeeld. In het proefschrift hebben we gekeken naar het meest interessante deelgebied van het Ising model: lage temperatuur en ferromagnetisch.

We hebben tot nu toe niet naar de omgeving gekeken. Als de energie van het systeem onafhankelijk is van zijn omgeving, zeggen we dat het systeem *vrije randvoorwaarden* heeft. Maar wat gebeurt er als de omgeving bestaat uit een ander materiaal met een vooraf gekozen spinconfiguratie? De waarden van de spins bij de rand voelen nu ook de invloed van de naaste buur spins van het andere materiaal.

In het algemeen bevat een stuk metaal een heleboel atomen. Een gram bevat al rond de  $10^{23}$  atomen. We willen graag naar volumes kijken van die orde van grote. We meten in aantallen atomen, dus ook de volumegrootte is in de orde van  $10^{23}$ . Eerst kijken we naar het Ising model met een groot eindig volume. Daarna proberen we om de uitdrukkingen hiervan te extrapoleren naar een 'oneindig' volume.

We laten het volume toenemen en voor elke stap nemen we de externe spinwaarden willekeurig  $+1$  of  $-1$ : *willekeurige randvoorwaarden*. Ook nemen we de temperatuur laag genoeg, zodat de spins sterk de neiging hebben om gelijke waarden aan te nemen. Het gedrag van de mate van overeenkomst van de spins blijkt af te hangen van hoe we het volume laten toenemen.

Uiteindelijk, door (A.66), zijn in alle voorkomende configuraties bijna alle spins hetzelfde. Omdat we de randvoorwaarden willekeurig hebben gekozen, heeft in de optredende configuraties desondanks de ene helft van de volumes bijna alle spins op en de andere helft bijna alle spins neer.

Als we diep in het volume ver weg van de rand kijken zien we nog steeds een effect van de randvoorwaarden. De lokale volume dichtheden van gebieden van gelijke spins zijn asymptotisch onafhankelijk van de randvoorwaarden. Maar zelfs voor hele grote volumes is er een significant effect op de spinwaarden. Als we kijken naar een heel groot volume, dan geldt met kans 1 dat ofwel alle configuraties bijna alle spins op hebben, ofwel bijna alle spins neer. De voorkeurswaarde is gelijk aan de spinwaarde die de meerderheid heeft van de externe spins van de randvoorwaarde met afstand 1 tot het volume. Omdat de temperatuur hoger is dan nul, heeft toch nog een klein deel van de spins een tegenovergestelde waarde.

Omdat we de volumes snel genoeg laten toenemen, zien we geen zogenaamde *mixtures*. Dit betekent dat het niet mogelijk is om met een positieve kans én configuraties te zien met de meeste spins op én configuraties met de meeste spins neer.



Deze onderwerpen hebben we behandeld in Hoofdstuk 4. Het was nodig om niet-triviale expansie technieken te introduceren: de *multi-scale* cluster-expansies. Onze multi-scale expansiemethode is geïnspireerd door ideeën van Fröhlich and Imbrie [35]. Het is een generalisatie van de meer bekende 'uniforme' cluster-expansie techniek. Met als doel vereenvoudiging van onze afschattingen kozen we ervoor om een andere representatie te gebruiken dan die in [35], en wel de zogenaamde *Kotecký-Preiss representatie*, die twee jaar later al werd ontwikkeld in [50].

Om bruikbare afschattingen te hebben, moesten we bepaalde *criteria* bewijzen: we hebben de convergentie nodig van sommige sommaties die gerelateerd zijn aan de expansies. Voor cluster-expansies is het cruciaal om te controleren of het Kotecký-Preiss criterium geldt. Voor onze expansies kan dit echter niet op een directe manier. Daarom introduceerden we een nieuw criterium en hebben we bewezen dat het equivalent hieraan is. Met dit nieuwe criterium kunnen we zelfs voor onze expansies bruikbare afschattingen verkrijgen. Het laatste hoofdstuk geeft meer uitleg over deze uniforme en multi-scale cluster-expansies.